CHAPTER SIX

RESULTS AND ANALYSIS: RURAL-URBAN LEVEL

6.1 Introduction

This chapter focuses on the difference of educational production function between rural and urban schools. The urban and rural schools used in this study have type of community differences stemming in part from their geographic locations. These distinct social environments are described to reveal how they indirectly shape the home educational resources of the students and their out-of-school study time, and influence student's academic achievement.

First, the general description of variables used is discussed. Then, the rest of this chapter evaluates how educational production process varies across rural and urban via Simple Regression Analysis and Canonical Correlation Analysis, reporting the result of Pearson Product-moment Correlation, Diagnostic Checking and Marginal Analysis. Finally, Section 6.9 gives the summary of the discussion in this chapter.

6.2 Description statistic

6.2.1 School outputs

As shown in Table 6.1, the mean achievement scores of rural schools are clearly below the average performance (50) for both mathematics and science. Conversely, the urban schools are achieving results which are above the average mean scores. This result is consistent with the findings from Leong et al. (1990) study, which has found that urban schools scored higher than rural schools in Malaysia. This is largely due to the majority of students in urban areas possess better home educational resource and more out-of-school study time as well as more experienced teacher. Comparing the school performance by their educational outputs, on average, students in urban schools scored slightly higher on mathematics whereas rural students were mastering in science.

The distribution of mathematics score for urban schools is right skewed, whereas it is left skewed for rural schools, as shown in Figure 6.1. The dispersion for schools in rural area is more dispersed than the dispersion for urban schools. The entire central 50% observations of the rural schools are below the median of urban schools. Comparing the mathematics score across school locations, there is significant difference as suggested by the independent samples t-test. Urban schools enjoy more upward mathematics score than rural schools. The distribution of science score is essentially similar with distribution of mathematics score. Using independent samples t-test, there is significant difference of science score across school locations.

Figure 6.1 Box-plots for school outputs by school location

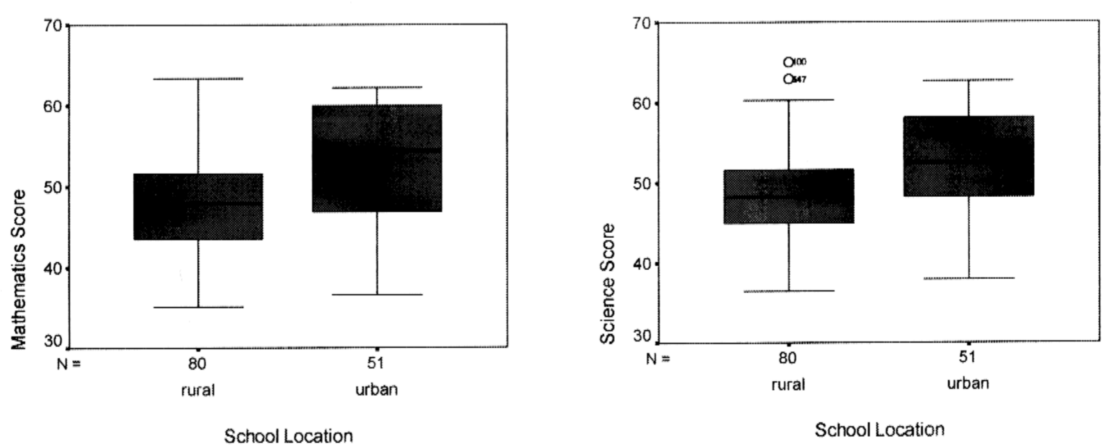


Table 6.1 Descriptive statistics for school outputs by school location

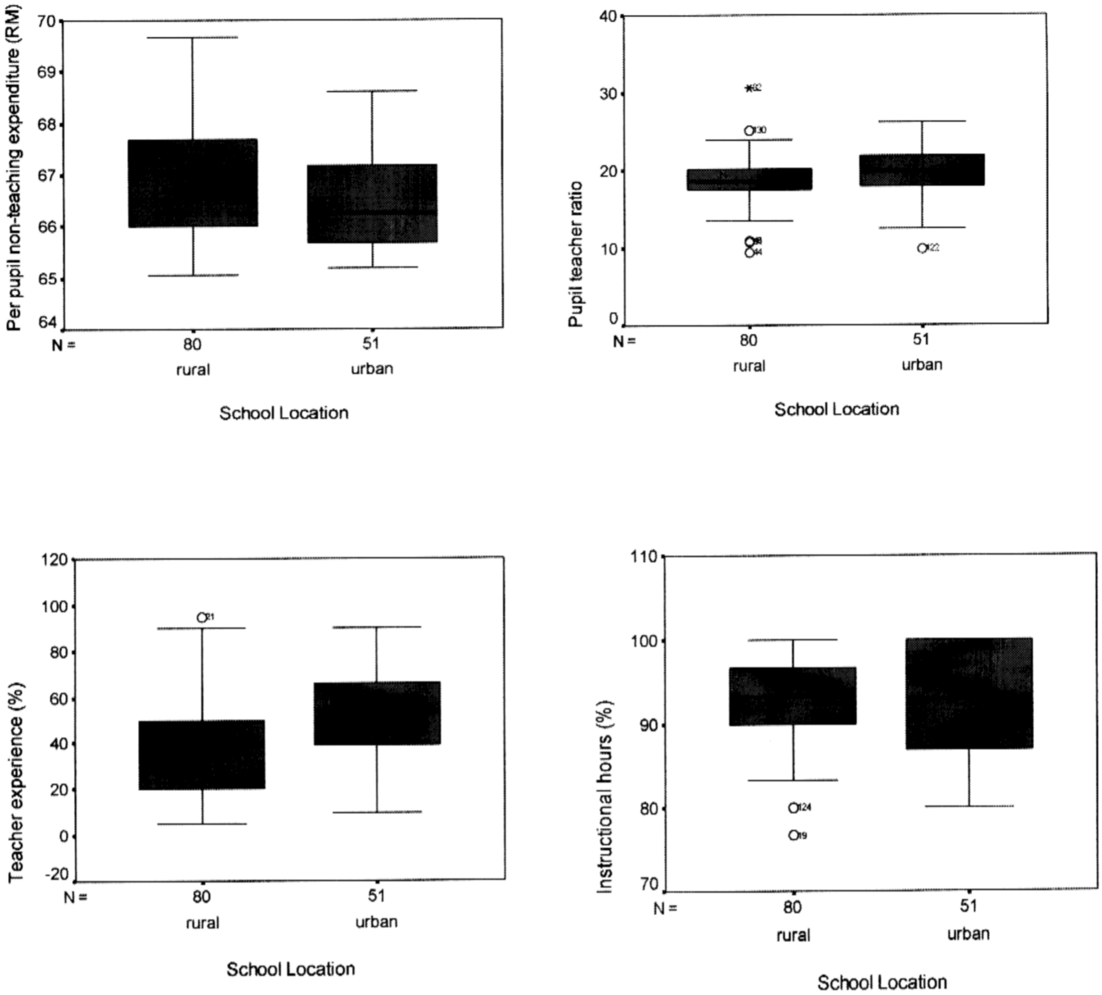
School Location	Mathematics Score		Science Score	
	Mean	Std. Dev.	Mean	Std. Dev.
Rural	48.115	6.489	48.563	5.871
Urban	53.307	7.294	52.283	6.937
	Statistics	p-value	Statistics	p-value
Levene test	2.054	0.154	4.169	0.430
t-test	-4.253	0.000	-3.293	0.001

6.2.2 School inputs

For the per pupil non-teaching expenditure, both distributions are left skewed with median taking the value of RM67. This is shown in Figure 6.2. The spread of the rural school's distribution is greater than the urban school's distribution. The difference of PPNTE between these two areas is significant, as presented in Table 6.2. However, we observe that the difference between the two means is very small, associate with a minor standard deviation.

The distributions of pupil teacher ratio for both locations look indistinguishable, as shown in the below box-plot. Both school locations are having symmetrical distributions; in addition, the dispersion is also similar. There is no significant difference of pupil teacher ratio between these two school locations.

Figure 6.2 Box-plots for school inputs by school location



For the percentage of teachers with more than five years experience, the distribution for rural schools is left skewed while the urban schools have right-skewed distribution. With comparison to urban schools, the rural schools having the widest spread of observation. Using pooled-variance t-test to confirm the difference between school locations, there is a significant difference.

The distributions for percentage of yearly school hours spent on instruction for each school location are right skewed. This is shown in Figure 6.2. The median for the rural schools is relatively higher than the urban schools. Using the separate-variance t-test to verify the difference in this variable between rural and urban schools, there is no significant difference.

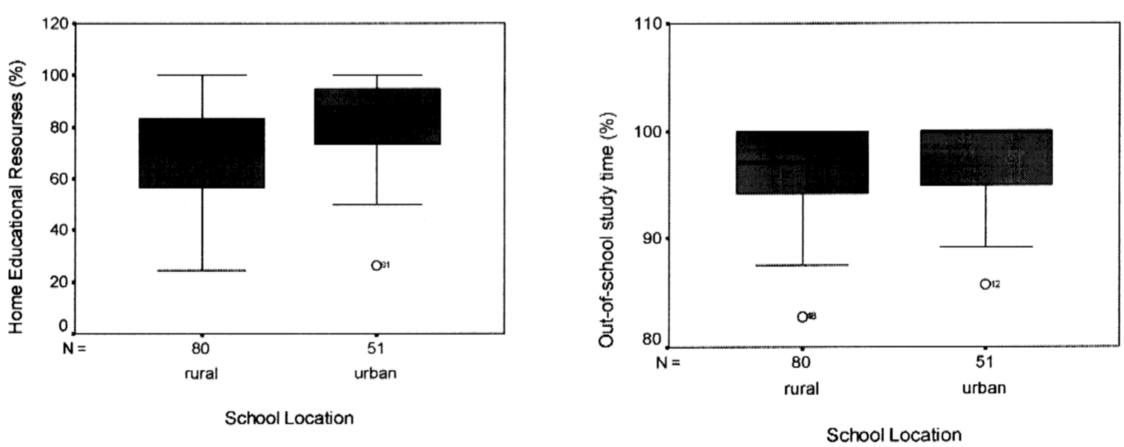
Table 6.2 Descriptive statistics for school inputs by school location

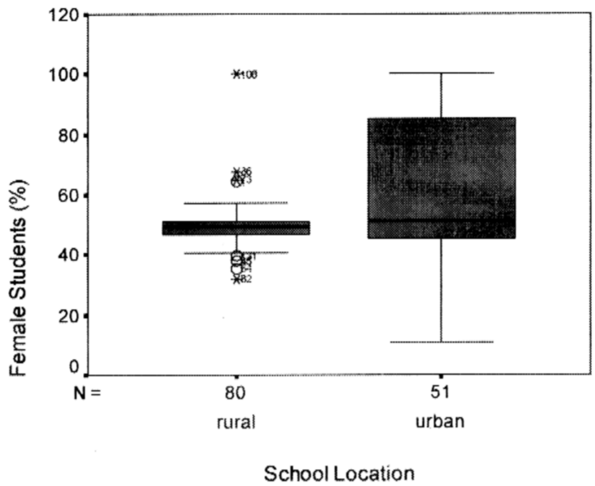
School Location	PPNTE		PTR		TE		INSHRS	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Rural	66.925	1.091	18.770	2.997	39.413	22.775	93.168	5.376
Urban	66.447	0.911	19.679	2.985	53.020	21.460	92.177	6.603
	Statistics	**p-value**	**Statistics**	**p-value**	**Statistics**	**p-value**	**Statistics**	**p-value**
Levene test	2.355	0.127	0.318	0.574	0.162	0.688	4.577	0.034
t-test	2.600	0.010	-1.694	0.093	-3.409	0.001	0.899	0.371

6.2.3 Environmental inputs

With regard to the percentage of students with at least medium level in home educational resources index, both distributions are right skewed, with the rural schools having the widest spread of observation. This is shown in Figure 6.3. The difference of HER is statistically significant, as show in Table 6.3.

Figure 6.3 Box-plots for environmental inputs by school location





The distributions of percentage of students with at least medium level in out-of-school study time index for both school locations are right skewed with a median of 97%. Generally, the distribution for rural schools is wider than the distribution for urban schools. The difference of OST between these two school locations is not statistically significant. The analysis revealed that urban schools with slightly less instructional hours and higher pupil teacher ratio, urban schools outperformed rural school because of its better teacher quality and environmental inputs.

Concerning the percentage of female students in class, the urban schools have right-skewed distribution whereas the rural schools have the left-skewed distribution. Clearly, the observation is least wide spread for rural schools. The gender difference between rural and urban is statistically significant.

Table 6.3 Descriptive statistics for environmental inputs by school location

School Location	HER		OST		FEMALE	
	Mean	SD	Mean	SD	Mean	SD
Rural	69.895	18.501	96.373	4.130	50.322	9.747
Urban	83.559	15.295	97.288	3.487	59.680	25.352
	Statistics	p-value	Statistics	p-value	Statistics	p-value
Levene test	2.770	0.098	2.155	0.145	67.195	0.000
t-test	-4.400	0.000	-1.312	0.192	-2.520	0.014

6.3 Simple Regression Analysis

In this section, we begin with a discussion on the simple regression analysis for mathematics achievement. As shown in Table 6.4, the test of goodness of fit (R^2) showing statistically weak results on all models with the exception for HER model in both rural and urban schools. This result is consistent with the finding at the national level.

The inputs variables which appear to be individually significant in both rural and urban schools are PPNTE, TE and HER. A comparatively higher R^2 is obtained in HER model for urban schools compares to rural schools. Undoubtedly, academic support in the home educational encourages learning more in the urban schools. The mathematics achievement increases by 0.34 scores for every additional 1 percent in the proportions of students with at least medium level in home educational resources index in urban schools compared to only 0.21 scores in rural schools.

The regression result concludes that the OST is statistically significant to explain the variation in the mathematics achievement only in urban schools with the highest positive coefficient 0.72. The FEMALE variable shown a smaller coefficient suggests a minor contribution on the urban school output production.

Table 6.4 Simple regression equation for explaining achievement in mathematics

The dependent variable : MATH

Variable	Rural				Urban			
	Coeff.	**SE**	**t-stat**	**p-value**	**Coeff.**	**SE**	**t-stat**	**p-value**
Constant	186.410	42.256	4.411	*0.000	239.565	71.219	3.364	*0.001
PPNTE	-2.066	0.631	-3.273	*0.002	-2.803	1.072	-2.616	*0.012
R^2 =	0.121				0.123			
RSS =	2924.606				2334.275			
F-stat =	10.714				6.841			
Constant	53.889	4.611	11.686	*0.000	46.781	6.883	6.797	*0.000
PTR	-0.308	0.243	-1.268	0.209	0.332	0.346	0.959	0.342
R^2 =	0.020				0.018			
RSS =	3259.176				2611.183			
F-stat =	1.607				0.919			
Constant	46.115	1.443	31.964	*0.000	49.259	2.703	18.227	*0.000
TE	0.051	0.032	1.599	**0.114	0.076	0.047	1.613	**0.113
R^2 =	0.032				0.050			
RSS =	3220.759				2525.971			
F-stat =	2.557				2.603			
Constant	49.013	12.754	3.843	*0.000	58.099	14.566	3.989	*0.000
INSHRS	-0.010	0.137	-0.071	0.944	-0.052	0.158	-0.330	0.743
R^2 =	0.000				0.002			
RSS =	3326.114				2654.276			
F-stat =	0.005				0.109			
Constant	33.555	2.309	14.532	*0.000	25.030	4.077	6.140	*0.000
HER	0.208	0.032	6.520	*0.000	0.338	0.048	7.050	*0.000
R^2 =	0.353				0.504			
RSS =	2152.940				1320.711			
F-stat =	42.511				49.695			
Constant	22.323	16.908	1.320	0.191	-16.695	27.392	-0.609	0.545
OST	0.268	0.175	1.527	0.131	0.720	0.281	2.557	*0.014
R^2 =	0.029				0.118			
RSS =	3229.794				2346.963			
F-stat =	2.331				6.539			
Constant	43.625	3.828	11.397	*0.000	48.823	2.569	19.006	*0.000
FEMALE	0.089	0.075	1.194	0.236	0.075	0.040	1.894	**0.064
R^2 =	0.018				0.068			
RSS =	3266.575				2478.735			
F-stat =	1.427				3.587			

Note: * Significant at 95% confidence level, ** Significant at 90% confidence level.

On the other hand, both OST and FEMALE variables are somehow individually statistically insignificant at 95% confidence level in rural schools. The negative sign for independent variable - PPNTE is the correct sign as expected in both school locations. Not surprisingly, the PTR and INSHRS are individually statistically not significant on explaining the variation on the mathematics performance at 95% confidence level for both school locations. Again, this is consistent with the finding at the national level.

Table 6.5 displays the simple regression analysis for explaining science achievement. Generally, the results obtained are similar with the results on explaining mathematics achievement partly due to the high correlation (0.9) between both achievement scores. Nevertherless, the OST variable is statistically significant on explaining science achievement at 90% confidence level with the highest positive coefficient 0.304 for rural schools and 0.911 for urban schools. But, the R^2 is relatively small to well-explain the variation in the science achievement scores. The intercept of 19.252 and -36.356 are statistically insignificant at any reasonable level of significance. Expectedly, the TE variable appears to be insignificant in both school locations.

Table 6.5 Simple regression equation for explaining achievement in science

The dependent variable : SCIENCE								
Variable	Rural				Urban			
	Coeff.	SE	t-stat	p-value	Coeff.	SE	t-stat	p-value
Constant	161.672	38.707	4.177	*0.000	184.404	69.798	2.642	*0.011
PPNTE	-1.690	0.578	-2.923	*0.005	-1.988	1.050	-1.893	**0.064
R^2 =	0.099				0.068			
RSS =	2454.014				2242.065			
F-stat =	8.541				3.584			
Constant	54.829	4.153	13.201	*0.000	46.131	6.547	7.046	*0.000
PTR	-0.334	0.219	-1.528	0.131	0.313	0.329	0.950	0.347
R^2 =	0.029				0.018			
RSS =	2643.645				2362.512			
F-stat =	2.334				0.903			
Constant	47.607	1.321	36.050	*0.000	48.594	2.576	18.865	*0.000
TE	0.024	0.029	0.834	0.407	0.070	0.045	1.543	0.129
R^2 =	0.009				0.046			
RSS =	2698.673				2294.545			
F-stat =	0.696				2.381			
Constant	41.974	11.515	3.645	*0.000	54.915	13.863	3.961	*0.000
INSHRS	0.071	0.123	0.573	0.568	-0.029	0.150	-0.190	0.850
R^2 =	0.004				0.001			
RSS =	2711.325				2404.268			
F-stat =	0.328				0.036			
Constant	36.980	2.214	16.699	*0.000	24.409	3.728	6.547	*0.000
HER	0.166	0.031	5.408	*0.000	0.334	0.044	7.599	*0.000
R^2 =	0.273				0.541			
RSS =	1980.150				1104.504			
F-stat =	29.251				57.741			
Constant	19.252	15.165	1.270	0.208	-36.356	24.671	-1.474	0.147
OST	0.304	0.157	1.935	**0.057	0.911	0.253	3.595	*0.001
R^2 =	0.046				0.209			
RSS =	2598.079				1903.869			
F-stat =	3.743				12.925			
Constant	44.338	3.460	12.813	*0.000	47.918	2.439	19.649	*0.000
FEMALE	0.084	0.068	1.243	0.218	0.073	0.038	1.942	**0.058
R^2 =	0.019				0.071			
RSS =	2669.840				2234.133			
F-stat =	1.546				3.770			

Note: * Significant at 95% confidence level, ** Significant at 90% confidence level.

6.4 Pearson Product-moment Correlation

6.4.1 Correlation between the independent variables

In Table 6.6, the Pearson correlation coefficients show low degree of correlation among all educational inputs.

Table 6.6 Pearson product-moment correlation coefficient matrix for school inputs and environmental inputs

	PPNTE	PTR	TE	INSHRS	HER	OST	FEMALE
			Rural Sample				
PPNTE	1.000						
PTR	-0.381**	1.000					
TE	-0.211	0.139	1.000				
INSHRS	0.202	0.053	-0.050	1.000			
HER	-0.212	-0.007	0.238*	-0.049	1.000		
OST	0.026	-0.280*	-0.077	-0.150	0.031	1.000	
FEMALE	-0.172	0.131	-0.042	-0.178	0.082	0.045	1.000
			Urban Sample				
PPNTE	1.000						
PTR	-0.545**	1.000					
TE	-0.068	-0.113	1.000				
INSHRS	0.127	0.064	0.207	1.000			
HER	-0.074	0.038	0.140	-0.045	1.000		
OST	-0.040	0.097	0.267	-0.063	0.258	1.000	
FEMALE	0.049	0.175	-0.133	0.154	0.335*	0.172	1.000

Note: * Significant at 95% confidence level, ** Significant at 99% confidence level.

6.4.2 Correlation between the dependent variables

As known, mathematics and science have a high degree of content overlap. Thus, the correlation between mathematics and science, 0.919 for rural schools and 0.942 for urban schools appear high enough to suggest that students who perform well in mathematics tend to do the same in science wherever he studies.

6.4.3 Correlation between the independent and dependent variables

Of the educational inputs, HER variable appears as the highest positive correlation with the school outputs. Comparing between school locations, the correlation is higher in urban schools than in rural school. Not surprisingly, the negative relationship between INSHRS and MATH happens in rural schools as well.

Table 6.7 Pearson product-moment correlation coefficient matrix for achievement scores, school inputs and environmental inputs

	PPNTE	PTR	TE	INSHRS	HER	OST	FEMALE
Rural sample: N = 80							
MATH	-0.348**	-0.142	0.118	-0.008	0.594**	0.170	0.134
SCIENCE	-0.314**	-0.170	0.094	0.065	0.522**	0.214	0.139
Urban sample: N = 51							
MATH	-0.350*	0.136	0.225	-0.047	0.710**	0.343*	0.261
SCIENCE	-0.261	0.135	0.215	-0.027	0.735**	0.457**	0.267

Note: * Significant at 95% confidence level, ** Significant at 99% confidence level.

6.5 *Canonical Correlation Analysis*

Table 6.8 represents the measures of overall model fit for canonical correlation analysis for both rural and urban schools. Statistically, canonical function 1 appears to be the preferable variate in both school locations. Besides that, the multivariate tests of significance for rural and urban schools are presented in Table 6.9, which indicate that the canonical functions, taken collectively, are statistically significant at the 0.05 level.

Table 6.8 Measures of overall model fit for Canonical Correlation Analysis

Canonical Function	Canonical Correlation	Canonical R^2	F-statistic	Probability
Rural Sample: N = 80				
1	0.696	0.485	4.997	0.000
2	0.358	0.129	1.770	0.118
Urban Sample: N = 51				
1	0.790	0.624	5.343	0.000
2	0.505	0.255	2.452	0.040

Table 6.9 Multivariate Test of Significance

Statistic	Value	Approximate F-statistics	Probability
		Rural Sample: N = 80	
Wilks'lambda	0.449	4.997	0.000
Pillai's trace	0.614	4.551	0.000
Hotelling's trace	1.089	5.446	0.000
Roy's ger	0.485		
		Urban Sample: N = 51	
Wilks'lambda	0.279	5.343	0.000
Pillai's trace	0.879	4.821	0.000
Hotelling's trace	2.005	5.872	0.000
Roy's ger	0.625		

The estimation results of the generalized production function for rural and urban schools using the Vinod procedure are given in Table 6.10, where we find the parameter estimates, canonical correlation and significance test information. In the urban schools, there is the higher canonical correlation coefficient of 0.79, which suggests a strong relationship between aggregate inputs and outputs. For both school locations, the Bartlett's χ^2 and Rao's F statistic indicate a rejection of the null hypothesis that all coefficients are equal to zero at any reasonable level of significance. Without doubt, this implied that the canonical estimates are significant overall, which support the hypothesis of jointness in production. The canonical redundancy analysis indicates that 60.17% of the total variance for the mathematics and science is explained by the educational inputs combination in urban schools, whereas only 46.39% in rural schools.

The parameter estimates have the expected sign for all the variables in both the two samples. However, both of the TE and INSHRS show a relatively small value for their parameter estimates in urban area; TE and FEMALE in rural schools.

Table 6.10 Canonical fit estimates of the Cobb-Douglas production function : full model

Variables	Parameter Estimates	
	Rural Sample	Urban Sample
School Outputs		
Mathematics Score	0.593	0.279
Science Score	0.428	0.734
School Inputs		
Per Pupil Non-teaching Expenditure (RM)	-0.436	-0.339
Pupil Teacher Ratio	-0.328	-0.148
Teacher Experience (%)	-0.056	0.036
Instructional Hours (%)	0.182	0.048
Environmental Inputs		
Home Educational Resources (%)	0.432	0.549
Out-of-school Study Time (%)	0.171	0.326
Female Students (%)	0.063	0.108
N	80	51
Rao's *F*	4.997	5.343
Wilk's Lambda	0.449	0.280
Bartlett's Chi-square	60.083	58.593
p value	0.000	0.000
R_c	0.696	0.790
Square R_c	0.485	0.624
R_d	46.39%	60.17%

Note: This presentation of canonical correlation analysis considers only the first canonical correlation. Standardized coefficients from the canonical correlation analysis are reported. The parameter are obtained in the solution of Equation (4.13)

6.6 Sensitivity Analysis

To investigate the importance of TE and INSHRS variables on determining the school performance in urban schools; and TE and FEMALE variables in rural schools, we re-estimated the equation 4.13 without the mentioned variables separately for both school locations. The results are display in Table 6.11 and Table 6.12.

In rural schools, dropping individually the TE variable and the FEMALE variable provides similar estimation results with the full model, which is shown in Table 6.11. Unexpectedly, the sign of parameter estimates of TE changed after deletion of FEMALA from the production process in rural schools. Thus, dropping simultaneously TE and FEMALE variables is carried out to examine effect of both variables in the production process of school outputs. It is surprisingly to reveal that the canonical redundancy index is remains stable and the parameter estimates are consistent with the full model. Therefore, we conclude that the TE and FEMALE variables do not add additional explanatory power to the education production function in rural schools.

On the other hand, the parameter estimates in Table 6.12 are remarkably stable and consistent with the model without TE and INSHRS for urban schools. The overall canonical correlations also remain stable. Hence, the analysis suggests that the TE and INSHR variables can be ignored from the full model for urban schools.

Table 6.11 Canonical fit estimates of the Cobb-Douglas production function : modified model for rural sample

Variables	Parameter estimates after deletion of		
	TE	FEMALE	TE & FEMALE
School Outputs			
Mathematics Score	0.657	0.599	**0.662**
Science Score	0.363	0.422	**0.358**
School Inputs			
Per Pupil Non-teaching Expenditure (RM)	-0.427	-0.441	**-0.432**
Pupil Teacher Ratio	-0.328	-0.319	**-0.318**
Teacher Experience (%)	Omitted	-0.055	**Omitted**
Instructional Hours (%)	0.179	0.170	**0.167**
Environmental Inputs			
Home Educational Resources (%)	0.426	0.433	**0.427**
Out-of-school Study Time (%)	0.170	0.178	**0.176**
Female Students (%)	0.062	Omitted	**Omitted**
N	80	80	**80**
Rao's *F*	5.487	5.847	**6.601**
Wilk's Lambda	0.471	0.452	**0.474**
Bartlett's Chi-square	56.860	59.936	**56.690**
p value	0.000	0.000	**0.000**
R_c	0.695	0.694	**0.692**
Square R_c	0.482	0.481	**0.479**
R_d	46.03%	46.04%	**45.68%**

Note: The final reduced model for rural sample (without the TE and FEMALE variables) is in bold type.

Table 6.12 Canonical fit estimates of the Cobb-Douglas production function : modified model for urban sample

Variables	Parameter estimates after deletion of		
	TE	INSHRS	TE & INSHRS
School Outputs			
Mathematics Score	0.244	0.314	**0.269**
Science Score	0.768	0.700	**0.744**
School Inputs			
Per Pupil Non-teaching Expenditure (RM)	-0.339	-0.335	**-0.331**
Pupil Teacher Ratio	-0.152	-0.142	**-0.146**
Teacher Experience (%)	Omitted	0.055	**Omitted**
Instructional Hours (%)	0.061	Omitted	**Omitted**
Environmental Inputs			
Home Educational Resources (%)	0.557	0.543	**0.553**
Out-of-school Study Time (%)	0.448	0.312	**0.332**
Female Students (%)	0.097	0.122	**0.109**
N	51	51	**51**
Rao's *F*	6.170	6.250	**7.468**
Wilk's Lambda	0.289	0.285	**0.293**
Bartlett's Chi-square	57.754	58.321	**57.761**
p value	0.000	0.000	**0.000**
R_c	0.790	0.789	**0.787**
Square R_c	0.624	0.623	**0.620**
R_d	59.96%	60.01%	**59.72%**

Note: The final reduced model for urban sample (without the TE and INSHRS variables) is in bold type.

As expected, the estimation result at rural-urban level reveals that the home educational resources effect is a strong influence on academic performance. Our analysis suggests that schools which have extra 1% in proportions of students with at least medium level of home educational resources index are likely to increase the academic performance by 0.55% and 0.43% respectively in urban and rural areas.

The negative coefficients of per pupil non-teaching expenditure seem to suggest there is inverse relationship between input and output. Similar result is obtained for both urban and rural areas. Comparisons were made to urban-rural difference; we found that the absolute values of parameter estimates for per pupil non-teaching expenditure variable are higher in rural schools. Obviously, rural area which associated with poor home educational resources depends heavily on the school inputs to increase performance academically.

Out-of-school study time variable appears to be positively affected the academic achievement more in urban schools than rural schools. This happened because majority of students in urban schools possess home environment conducive to educational development. Perhaps, rural areas may not stress education as much as urban areas. Generally, students need to spend the extra time to keep up academically in order to increase the school performance.

As expected, pupil teacher ratio has expected negative coefficients. Comparison was made that the relationship of performance to pupil teacher ratio has been found to be disappointingly weak in urban schools. This is consistent with the finding by Kiesling (1967).

With regard to the yearly school hours spent on instruction, the analysis in Table 6.11 shows that instructional hours have a positive effect on aggregate output in rural schools. Nevertheless, the impact of instructional hours is greater in rural schools than urban schools.

In urban schools, percentage of female students in class explains a small variation in achievement even when other school and environmental factors are controlled. On the other hand, gender variable does not matter performance in rural schools.

6.7 Diagnostic Checking

In this section, we present the results of testing the assumptions of multivariate analysis.

6.7.1 Normality

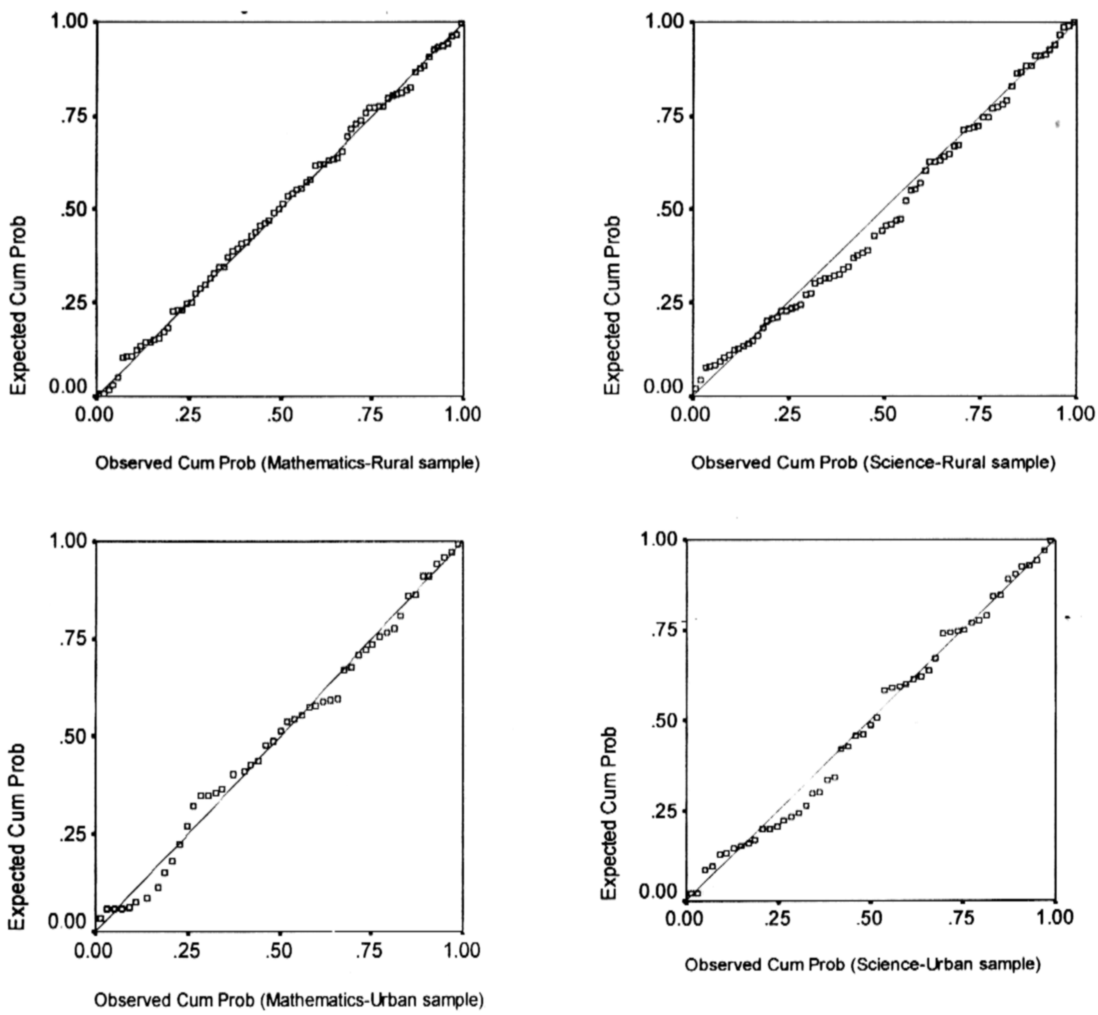
As shown on the normal probability plots in Figure 6.4, the standardized residuals can be considered scattering around the diagonal line closely in both school locations. These findings are support by the analysis of Kolmogorov-Smirnov Z in Table 6.13.

In table 6.13, the Komologorov-Simirnov Z indicates a probability of 0.998 and 0.733 for standardized residuals of mathematics and science for rural schools; 0.947 and 0.959 for standardized residuals of mathematics and science for urban schools. These large significance values of one-sample Kolmogorov-Simirnov test indicate that both standardized residuals of mathematics and science do not differ significantly from normal distribution for both school locations.

Table 6.13 One-sample Kolmogorov-Smirnov test comparing distribution of standard residuals of school output

Test Variable	N	Normal Parameters		Most Extreme Differences			K-S Z	2-tail Sig.
		Mean	Std. Dev.	Absolute	Positive	Negative		
Rural sample								
MATH	80	0.000	0.967	0.043	0.040	-0.043	0.387	0.998
SCIENCE	80	0.000	0.968	0.077	0.077	-0.054	0.687	0.733
Urban sample								
MATH	51	0.000	0.950	0.073	0.071	-0.073	0.524	0.947
SCIENCE	51	0.000	0.947	0.071	0.071	-0.054	0.507	0.959

Figure 6.4 Normal P-P plots of standard residuals for school outputs

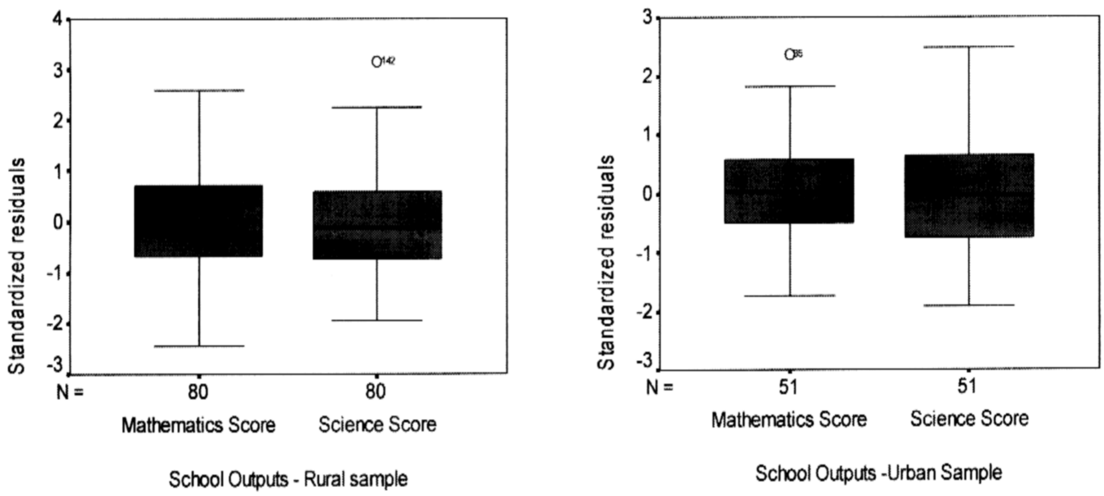


Note: According to Ruggiero (1996, 1998), the canonical regression technique still performs well even violating normality. Refer to Chapter 4 for more details.

6.7.2 Outliers

Figure 6.5 illustrating only one outlier in the distribution of both mathematics and science standardized residuals and suggests that the data is with no obvious outliers. Therefore, our estimation at rural-urban level is not much influenced by outlier problem.

Figure 6.5 Box-plots for standardized residuals of school outputs



6.7.3 Multicollinearity

The *F*-test results for significant of Auxiliary Regressions with 95% confidence level were reported in Table 6.14. The results indicate that two out of five independent variables in rural schools and three out of five independent variables in urban schools with $F_i > F_{0.05;\,4,\,126} = 2.42$. This meant that the problem of multicollinearity, however, is not serious. However, Variance Inflation Factor (VIF) as an indicator of multicollinearity noted the different results as Auxiliary Regression, where the VIF for all the variables do not exceed 2.5, which is much below the benchmark level of 10. So, the problem of multicollinearity is considered mild in these model and the educational inputs in the models are not highly correlated for rural and urban schools.

Table 6.14 The significant of auxiliary regression for testing the presence of multicollinearity

Dependent Variable	Independent Variables	R_i^2	F_i-statistic	$VIF = \frac{1}{1-R_i^2}$
	Rural Sample			
PPNTE	PTR, INSHRS, HER, OST	0.240	5.828	1.316
PTR	PPNTE, INSHRS, HER, OST	0.240	5.835	1.316
INSHRS	PPNTE, PTR, HER, OST	0.068	1.351	1.073
HER	PPNTE, PTR, INSHRS, OST	0.051	0.997	1.054
OST	PPNTE, PTR, INSHRS, HER	0.100	2.048	1.111
	Urban Sample			
PPNTE	PTR, HER, OST, FEMALE	0.330	5.675	1.493
PTR	PPNTE, HER, OST, FEMALE	0.347	6.104	1.531
HER	PPNTE, PTR, OST, FEMALE	0.169	2.344	1.203
OST	PPNTE, PTR, HER, FEMALE	0.080	1.001	1.087
FEMALE	PPNTE, PTR, HER, OST, FEMALE	0.109	2.534	1.122

6.7.4 Linearity

The graphical depictions of the relationship between the two canonical variates (aggregate output and aggregate input) for rural and urban schools are given in Figure 6.6 and 6.7 respectively. As expected, the plots demonstrate a strong linear relationship between the two composite variables.

Figure 6.6 Scatter plot of canonical variables: Rural sample

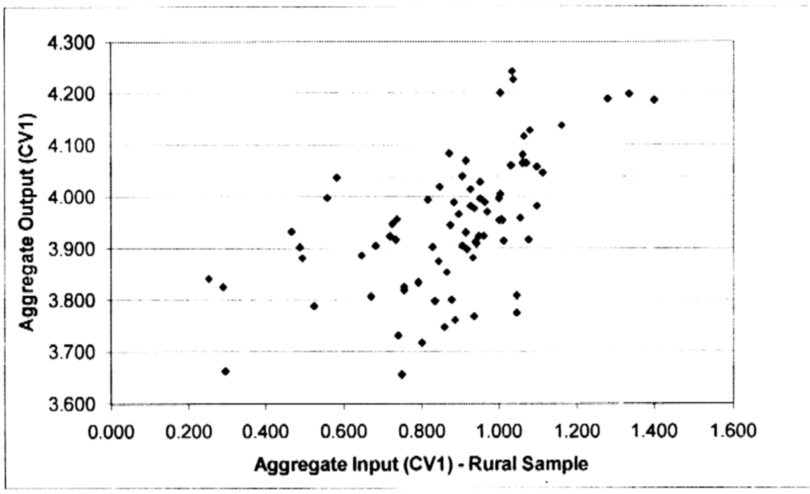
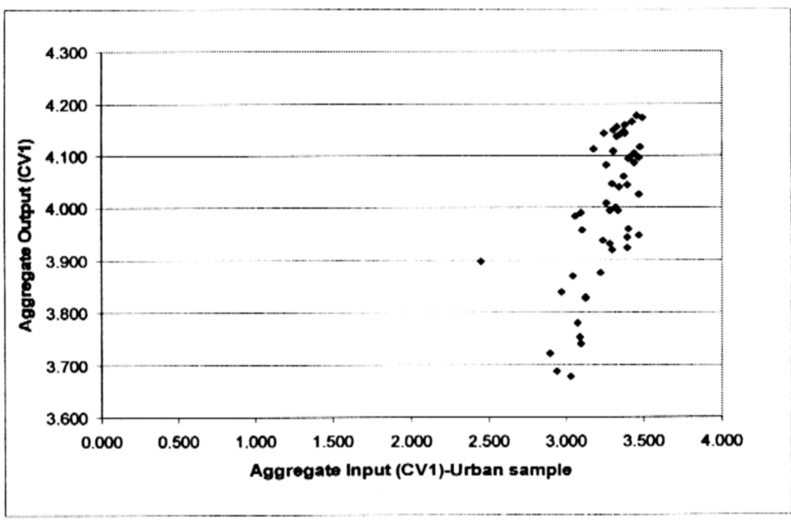


Figure 6.7 Scatter plot of canonical variables: Urban sample



6.8 The Marginal Analysis

Using the parameter estimates in Table 6.11 (final reduced model for rural schools - without TE and FEMALE model) and Table 6.12 (final reduced model for urban schools - without TE and INSHRS model), the marginal rate of output transformation showing the relationship among outputs and the marginal elasticity linking the input and output can be calculated. These calculations are displayed in Table 6.15.

The marginal rate of transformation between the two outputs, calculated at the means of the outputs is -0.536 and -2.820 for rural and urban schools respectively. Clearly, the MRT for both rural and urban schools has show the difference in their productivity of inputs on the production of outputs. For urban schools, MRT of mathematics for science greater than 1 implying that the input productivities in mathematics fall more, when the productivities in science increase, whereas, MRT less than 1 means that very little mathematics production is lost to obtain 1 extra score of science in rural schools.

Table 6.15 Estimated relationships between inputs and outputs

Variables	Rural Sample		Urban Sample	
	Math	Science	Math	Science
Marginal Rate of output transformation				
School Outputs				
Math Score	-1.000	-1.866	-1.000	-0.355
Science Score	-0.536	-1.000	-2.820	-1.000
Marginal Elasticity				
School Inputs				
Per Pupil Non-teaching Expenditure (RM)	-0.652	-1.206	-1.232	-0.445
Pupil Teacher Ratio	-0.481	-0.889	-0.541	-0.196
Instructional Hours (%)	0.252	0.466	NR	NR
Environmental Inputs				
Home Educational Resources (%)	0.645	1.195	2.054	0.743
Out-of-school Study Time (%)	0.266	0.491	1.235	0.446
Female Students (%)	NR	NR	0.407	0.147

Note: 1. Marginal rates of output transformation and marginal elasticity are calculated by using Equation (4.17) and (4.51) using parameter estimates from Table 6.11 and 6.12. Marginal rates of output transformation are calculated at the means of the respective outputs.

2. NR=Not related (omitted variable)

A 1% increase in the percentage of students with at least medium level of home educational resources index would leads to a 0.65% increase in mathematics scores and 1.20% increase in science scores for rural schools. Whereas, in urban schools, it could leads to 2.05% and 0.74% increase in mathematics and science scores, respectively. There is no specific pattern of marginal effect of educational inputs on outputs in urban and rural schools, which differs depending on the outputs. Generally, the marginal effect of educational inputs on mathematics is higher than science in urban schools. While, rural schools have higher marginal effect on science than mathematics.

Table 6.16 Marginal Product

Variables	Rural Sample		Urban Sample	
	Math	Science	Math	Science
Marginal Product				
School Inputs				
Per Pupil Non-teaching Expenditure (RM)	-0.469	-0.875	-0.988	-0.350
Pupil Teacher Ratio	-1.233	-2.301	-1.466	-0.520
Instruction Hours (%)	0.130	0.243	NR	NR
Environmental Inputs				
Home Educational Resources (%)	0.444	0.829	1.310	0.465
Out-of-school Study Time (%)	0.133	0.247	0.677	0.240
Female Students (%)	NR	NR	0.363	0.129

Note: NR=Not related (omitted variable)

Table 6.16 concludes the marginal productivities of each input with respect to mathematics and science. Marginal product of science achievement with respect to all educational inputs is relatively higher than mathematics achievement in rural schools. The reverse is true for urban schools where its marginal product of mathematics achievement with respect to all educational inputs succeed its science achievement.

Generally, economic opportunities are more abundant in the urban area than in the rural. Youths in urban are explore to the commercial environment sharpen their mathematical thinking. In sharp contrast, in the area surrounding rural school, most local employment is directly related to agriculture. Many jobs entail working "in the fields". Parents of students sometimes work picking fruit or working machinery in the agricultural fields that surround the town. Thus, our result suggests that urban schools are more productive in the production of mathematics, whereas rural schools in science.

Table 6.17 summarizes the marginal rates of substitution for rural and urban schools. For the rural schools, the MRTS of INSHRS for OST (which equals MP_{INSHRS}/MP_{OST}) has a value of 0.981. This result reflects that rural schools with majority low level out-of-school study time' students can compensate for deficiencies in this environmental factors with extra instructional hours of 1.02% is necessary to trade-off a 1% of disadvantage in percentage of students with at least medium level in out-of-school study time index (OST) and keep output constant. Thus, a 1 standard deviation (SD) decrease in OST can be offset by 1.3 SD increase in instructional hours.

Since INSHRS variable does not determine the production of urban schools' outputs, implicitly, it shows that it is easier for rural schools to compensate than it is for urban schools. Empirically, this analysis appears to support the effectiveness of instructional hours through extra tuition to improve the school performance in rural areas. Consequently, investments in improved instructional hours are likely to have encouraged effect on educational outputs in rural schools. This result seems to support the recent RM200 million tuition voucher scheme for primary school-children from poor families in rural schools and certain poor urban schools to obtain extra lessons. In further extent, we foresee that this assistance would help the lower secondary school-children as well because the lower secondary education, which more emphasis on basis cognitive skills than in the later grades needs extra lessons as the primary level.

Table 6.17 Marginal Rates of substitution derived from Cobb-Douglas production function estimates of Table 6.11, 6.12, 6.15 and 6.16

	PPNTE	PTR	INSHRS	HER	OST	FEMALE
			Rural sample : N = 80			
PPNTE	1.000	2.628	-0.277	-0.947	-0.283	NR
PTR	0.381	1.000	-0.106	-0.360	-0.108	NR
INSHRS	-3.605	-9.474	1.000	3.413	1.019	NR
HER	-1.056	-2.776	0.293	1.000	0.299	NR
OST	-3.538	-9.299	0.981	3.349	1.000	NR
			Urban sample : N = 51			
PPNTE	1.000	1.484	NR	-1.326	-0.685	-0.368
PTR	0.674	1.000	NR	-0.894	-0.461	-0.248
HER	-0.754	-1.119	NR	1.000	0.516	0.277
OST	-1.461	-2.167	NR	1.937	1.000	0.537
FEMALE	-2.720	-4.036	NR	3.607	1.862	1.000

Note: 1. Property of joint production functions suggests that the MRTS between any two inputs are the same for either output. Therefore, we only report a single matrix each for urban and rural.

2. NR = Not related (omitted variable)

6.9 Summary of Main Findings

This chapter has presented the result of rural-urban disparities in school academic achievement. The results obtained relevant to the difference between rural and urban mathematics and science performance, which have been briefly summarized and presented in tabular form. The very brief narrative highlighted the most important results. The tables presented have also very clearly highlighted the pertinent results.

School qualities vary across rural and urban. We conclude that instructional hour is an effective and important variable that could be used to offset the disadvantage in low level of out-of-school study time to the educational production function in rural areas. Besides that, this result suggests that urban schools are more productive in the production of mathematics and rural schools in science. This analysis demonstrates the significant effect of home educational resources on Malaysian schools' mathematics and science achievement. As expected, teacher experience appears to have the smallest impact on the educational outputs in both rural and urban schools, in which the data suggests to remove it from the full model.